



专题：数据网与算力网络

Klonet: 面向技术创新的网络模拟实验平台

谢景昭¹, 单炜¹, 肖畅¹, 马铁¹, 陈力², 虞红芳¹, 孙罡¹

(1. 电子科技大学, 四川 成都 611731; 2. 华为技术有限公司, 广东 深圳 518129)

摘 要: 网络关键技术的创新是推动网络未来发展和整个行业进步的使能条件, 为此亟须一个能对创新技术进行方便、快捷验证的网络实验平台, 以有效降低创新门槛、促进创新趋势。通过分析网络模拟技术在网络实验中的优势, 结合网络技术创新对网络实验平台的需求, 提出了面向技术创新的网络模拟实验平台 Klonet。该平台可以灵活地扩展规模、全方位地对实验网络进行模拟、细粒度地管理模拟网络, 并促进实验的工作流程。结合两个具体的用例, 分析了在 Klonet 上进行实验的方法, 验证了 Klonet 支持网络技术创新的能力。

关键词: 网络实验; 网络模拟; 网络虚拟化

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2021241

Klonet: a network emulation platform for the technology innovation

XIE Jingzhao¹, SHAN Wei¹, XIAO Chang¹, MA Tie¹, CHEN Li², YU Hongfang¹, SUN Gang¹

1. University of Electronic Science and Technology of China, Chengdu 611731, China

2. Huawei Technologies Co., Ltd., Shenzhen 518129, China

Abstract: The innovations of key technologies enable and promote the future development of networks and the network industry. Therefore, there is an urgent need for a network experimental platform that can easily and quickly verify innovative network technologies, so as to effectively lower the innovation barriers and promote innovations. By analyzing the advantages of network emulation technology in network experiments and combining the demand of network technology innovation for network experiment platform, a network emulation experiment platform Klonet for network technology innovation was proposed. The platform can flexibly expand the scale of the network, emulate the network in all perspectives, manage the emulated network in a fine-grained manner, and thus ease the workflow of network experimentation. The experimental method on Klonet was described and the ability of Klonet to support network technology innovation was verified based on two specific use cases.

Key words: network experiment, network emulation, network virtualization

收稿日期: 2021-09-18; 修回日期: 2021-10-14

通信作者: 虞红芳, yuhf@uestc.edu.cn

基金项目: 国家重点研发计划项目 (No.2019YFB1802800); 华为资助项目 (No.YBN2019115092)

Foundation Items: The National Key Research and Development Program of China (No.2019YFB1802800), HUAWEI Research Foundation (No. YBN2019115092)

1 引言

工业互联网、远程医疗、VR/AR 等新兴应用场景与 5G 通信、云网融合等新型承载技术的出现,对数据网的内生能力和性能指标提出了更高要求。网络技术创新作为数据网发展的根本动力,是应用和业务创新的使能条件,是关乎整个行业进步的关键环节。为了凝聚共识,以顶层设计的方式提出面向未来的网络创新目标,推进关键技术创新和行业发展,也有网络 5.0 这样的网络跨代演进概念被提出。然而创新技术的开发和验证仍然面临许多现实性的难题,如成本、效率、安全以及系统复杂性和可重复性等,严重阻碍了技术的创新,因此亟须一个网络创新平台为其提供可用且友好的开发和实验环境。可见网络创新平台对网络技术的创新具有重要意义,是网络未来发展的关键。

网络创新平台提供新技术的开发环境以及用于评估和验证创新技术的具体实验场景,通常包括网络仿真平台和网络模拟平台两类。与完全基于软件模型的仿真平台相比,网络模拟平台依赖实体的模拟网络进行技术的开发和实验,因而更接近实际的网络运行、更具有现实意义,模拟平台上开发的网络新技术也可以更快速地落地部署。尽管目前的网络模拟平台能在一定程度上支持技术创新,但大多仍存在规模有限、适用的创新形式单一、可执行的操作少、实验开展难度高等问题,没有充分挖掘模拟平台的优势,难以应对网络未来发展过程中规模大且场景复杂、技术创新多样化且发展迅速的趋势。

针对网络未来的发展需要,本文提出了一种面向技术创新的网络模拟实验平台 Klonet。首先,Klonet 模拟的网络可以根据需求灵活扩展,支持网络层、传输层、应用层等多层次网络技术的创新,符合网络进入超大规模时代的特征以及从顶层设计网络的目标;其次,Klonet 支持在实验的

运行过程中引入动态网络事件,并为网络实验的完整流程提供相应的辅助功能,使得实验场景更加真实、过程更加便捷,加速了创新概念转化为实际部署的进程,从而推动网络技术的创新发展。通过两个具体用例分析说明了使用 Klonet 进行网络创新的方法,并且展现了 Klonet 为技术创新带来的多种可能性。

2 网络创新平台的研究现状

通过对比网络创新平台两类方法的不同实现思路,阐述网络仿真的局限性以及网络模拟在开发和测试网络新技术上的优势,并对几种具有代表性的网络模拟平台进行分析,同时指出它们面对技术创新时的不足。

当前存在多种用于开发、评估和验证新技术的网络创新平台,按照本质实现方法的不同将这些平台分为两类:网络仿真平台和网络模拟平台。网络仿真平台依据离散事件模型为真实网络的逻辑结构和运行状态建模,通过逐层嵌套的软件模块构建网络系统抽象,是一种完全基于软件的实现方法,比较具有代表性的仿真平台有 OMNeT++^[1]、NS3^[2]和 OPNET^[3]等。网络模拟平台则通过物理设备或虚拟化技术构建实际的网络,对真实网络进行模拟,该模拟网络与目标网络的原型相同或是对其最小规模的近似,是模拟平台的实现基础,其中比较有代表性的模拟平台包括 PlanetLab^[4]、Emulab^[5]和 Mininet^[6]等。

2.1 网络模拟与网络仿真对比

网络仿真依赖于对网络的建模,是一种在事件、时序、语义层面对网络的模仿,所以网络仿真过程中没有真实的网络应用、协议或报文,因此网络仿真存在许多固有的局限性。

首先,仿真模型要尽量完美地反映、捕捉现实中的实体对象、操作逻辑和动作行为,因此对建模的要求很高、建立模型的过程很难;其次,仿真只是一种逻辑和数学上的表示和计算,而不



是网络系统中的真实实现，因此通过仿真实现和测试的网络创新与能够在实际网络中部署的成品间还存在非常大的差距；最后，仿真实验中的理想模型与真实网络环境中的实际运行情况也很难保持完全一致，因此对网络运行的还原度较低。可见依赖网络仿真方法构建的创新环境，只能对网络创新做相对理想化的分析和评估。

网络模拟通过物理或虚拟形式的模拟网络来分析真实网络系统中的行为和表现，以此对运行在模拟网络上的创新技术进行评估和验证。模拟网络在终端节点方面，包含真实的操作系统，具有完整的内核协议栈，可发送标准分组，可运行真实的网络应用；在网络节点方面，运行真实网络设备的固件或软件，执行真实的网络协议，可交换和转发标准分组；在网络工具方面，具有同真实网络一致的测试工具、配置工具和程序脚本等，其交互模式同生产网络一致。

对比基于软件模型的网络仿真，模拟方法中的各类对象和事件是真实存在和发生的，不依赖烦琐的模型语言描述；在模拟网络上开发和测试创新实例，所使用的方法和工具与在实际网络中相同，创新成果更接近于能够在现网中部署的形式；模拟方法中的测试结果是网络实体间相互作用的真实体现，因此得到的评估和验证结果也更能反映网络创新在真实运行情况下的性能和表现。总而言之，网络模拟能够高度还原真实网络的运行过程，基于模拟开发的网络创新实例更接近于实际部署，实验结果更为可信。

2.2 现有网络模拟平台

许多机构和高校设计实现了网络模拟平台用于网络研究和技术创新工作，其中常见的有如下3个。

(1) SDN 原型测试工具 Mininet

Mininet 支持在单台主机上快速地构建大型 SDN，凭借其轻量、易用等特点 Mininet 已经成为近几年热门的网络模拟工具。利用 Linux 的容器技

术，Mininet 可以轻松地在单台主机上快速搭建一个包含上百网络节点和吉比特级别链路的虚拟网络，以支撑 SDN 系统的创新和研究工作。然而，受限于单机的计算资源，Mininet 不能支持大规模或资源敏感的网络技术、应用进行创新和实验。

(2) 分布式系统开发平台 PlanetLab

PlanetLab 向网络研究人员提供全球覆盖的网络平台，以支持全球级别大规模网络服务或应用的开发工作^[7]，自上线以来取得了非常大的成功。PlanetLab 通过虚拟机向用户提供上层的 overlay 网络，服务切片初始化后可作为独立的网络实验资源被用户访问和控制，用于业务和应用的创新开发。然而，PlanetLab 只模拟端到端的应用层网络，缺乏对网络拓扑和网元功能的考虑，无法用于底层网络技术的创新。

(3) 网络控制平面模拟器 CrystalNet

CrystalNet^[8]是微软提出的网络模拟平台，利用 Docker 模拟大规模的生产网络，并承诺其具有高保真的控制平面。新网络架构、设备固件和配置策略被实际部署之前，可以首先利用 CrystalNet 进行测试以验证其正确性，从而减少各种未知错误可能对生产网络造成的影响以及随之而来的经济损失。然而，CrystalNet 仅模拟网络的控制面，不提供网络的数据面属性，难以应用于与数据面相关的技术创新之中。

综上所述，现有的网络模拟平台常常聚焦于网络的个别特征，缺乏全方位模拟大规模网络的能力，只支持特定类型或特定场景下的网络技术创新。并且这些平台部署的模拟网络大部分是“静态的”，缺少细粒度的网元、链路操作能力，很难模拟像链路增加、移除这样的网络动态事件。此外很少有平台考虑专门辅助网络实验的功能，因此需要额外借助多种工具用于分析研究，不利于新技术的评估和验证。可见，如何设计一个以技术创新为目标的网络模拟实验平台，仍是一个充满挑战且十分重要问题。

3 Klonet 的设计目标

鉴于已有平台在功能特性上的缺陷和不足,以网络发展现状、技术创新的一般性要求以及网络实验的常见需求为基础,分析了 Klonet 的总体设计目标。

(1) 灵活的规模可扩展能力

不断扩大的网络规模是通信网络发展的显著特征之一,并且随着网络向工业、医疗、交通和市政等领域的垂直延伸,仍有更多的人和海量的新设备在持续接入互联网络。这意味着许多未来的网络难题和网络创新理念是以大规模的网络系统、网络应用为背景的,也意味着网络创新技术的开发和实验需要大规模的网络环境作为支撑。因此, Klonet 要灵活应对网络技术对实验规模的需求,其部署的模拟网络,以及相应的平台资源和平台整体架构都应该易于规模的扩展。

(2) 全方位的网络模拟能力

某些网络技术依赖网络的部分特征就能进行开发和实验。例如,分布式应用只重视端到端链路的数据面属性,网络拓扑、网络节点和网络操作并不受重视。但是就路由协议的创新而言,又需要关心拓扑结构和节点行为,仅模拟端到端链路明显不够。所以要真正服务于网络的顶层设计,满足网络技术从 3 层到 5 层的创新需求,网络的拓扑结构、网络层次、网元功能、控制面行为以及数据面属性等因素都不可或缺。因此, Klonet 要对实际网络进行全方位的模拟。

(3) 细粒度的网络管理能力

由于在网络的运行过程中时常有动态的事件发生,所以网络状况并不是一成不变的,而相关网络技术对这些事件的响应能力则很可能是其性能评估的重点。例如,当有新的终端设备加入同条链路时,公平分配链路带宽的能力是拥塞控制算法的关键特性。为了使新技术在更真实的复杂环境下进行开发和实验,网络创新平台应该对动

态的网络场景进行模拟,这意味着平台的模拟网络应该支持动态的调整。因此, Klonet 要支持实验过程中对网元和链路等网络元素的细粒度管理操作。

(4) 实验工作流的支撑能力

开展网络实验是复杂的过程。完整的实验工作流主要包括网络拓扑设计、网络参数配置、网络流量生成以及最后的实验数据获取等步骤^[9],该流程中的大部分工作很烦琐且容易出错,并且在模拟网络中生成网络流量、收集网络运行数据还需额外借助多种网络应用和工具。所以当实验规模较大时,评估和验证新技术的难度以及花费的大量时间都显而易见,这极大阻碍了创新概念的快速落地。因此,为了加快技术创新的进程, Klonet 要提供与某些实验步骤相关的支撑功能。

4 Klonet 技术方案

针对上述设计目标,从平台的系统架构、模拟网络的构建方法、模拟网络管理的关键实现以及支撑实验的具体功能 4 个方面,对本文面向技术创新的网络模拟实验平台 Klonet 展开论述。

4.1 系统架构

随着网络规模的不断扩大,规模扩展能力将会成为网络创新平台最重要的特性。系统架构作为平台的根本设计,关乎基础资源提供、基础网络性能保证以及模拟网络扩展等重要问题。Klonet 系统架构如图 1 所示。其中,基础设施是平台网络和计算资源的根本;模拟网络是平台部署的虚拟网络,是开展网络技术创新的实际环境;控制框架是平台各类控制逻辑的集合,负责解析指令并执行各类操作;应用接口是用户与平台交互、开展实验的接口。

由于资源限制,单台主机无法承载大规模的模拟网络,更不用谈有效的性能保障。所以 Klonet 的基础设施由多台高性能物理主机和其底层的高性能物理网络(高带宽、低时延)共同组成。一

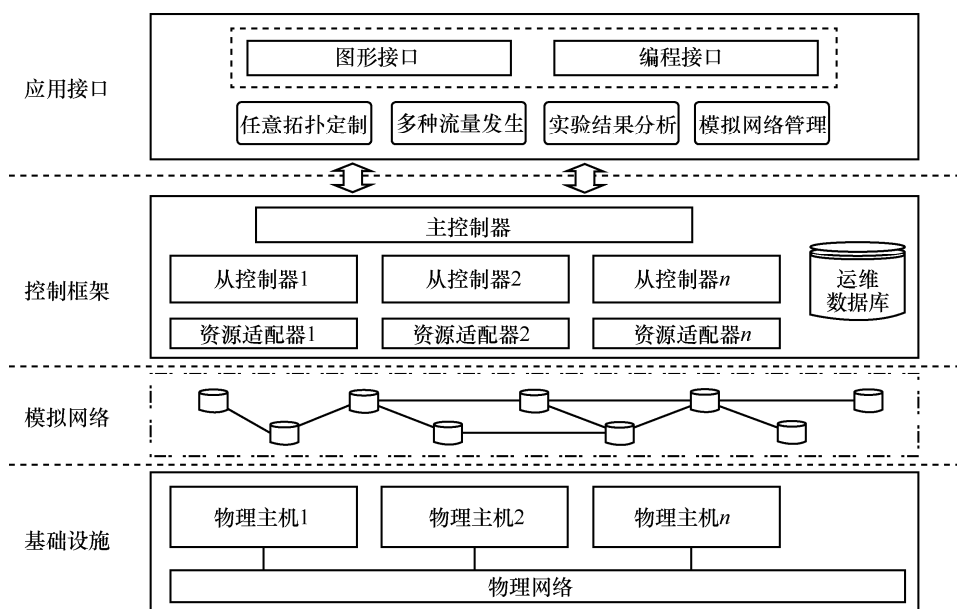


图1 Klonet 系统架构

方面充足的计算能力可以满足实验规模扩展的资源需求，同时还能减小干扰、保证性能隔离；另一方面在一定范围内可信且高性能的物理链路，能够为模拟网络链路属性的模拟提供基本的网络性能支持。

技术创新环境可以灵活扩展，意味着模拟网络能够在底层分布式的基础设施上任意地扩展部署，所以 Klonet 通过计算和网络虚拟化技术，利用虚拟网元和虚拟链路，在物理 underlay 网络的基础上，搭建统一的 overlay 网络用于网络创新实验，解决了扩展创新实验规模的首要问题。

当规模扩展时，控制框架必须将底层分布式的资源、分散部署的模拟网络以及上层集中的用户接口有效地整合到一起。为此 Klonet 采用分层的控制框架，通过程序与数据的分离、主从控制逻辑的拆分以及动作执行和软硬件接口的解耦，在有效整合上下层的同时进一步提高平台的可扩展性。首先，控制框架将程序处理所需的关键数据从进程内存中剥离，通过统一的运维数据库维护和存储，从可存储的数据规模方面提高了可扩展性，同时也提高了平台运行的可靠性。其次，

控制架构按照主从的形式拆分控制逻辑。主控制器是平台控制逻辑的核心，对上层的用户指令进行分析和处理，并向下层的各个从控制器发送相应的控制信号；从控制器则与物理主机一一对应，根据自身逻辑分析下达的命令，对主机执行相应的操作。主从设计在降低集中控制逻辑处理复杂度的同时，通过一定程度上的自治能力降低了多主机管理的难度，更利于规模的扩展。最后，Klonet 凭借资源适配器将主机的各类软硬件资源封装为统一的网元操作和链路操作接口（如网元的创建、停止、销毁等）供从控制器调用，使逻辑上的执行与具体的资源操作解耦，不仅降低了从控制器管控多维资源的难度，更提高了控制框架对底层资源的兼容性和平台规模的扩展能力。

面对底层分布式的物理资源以及分散部署的模拟网络，用户需要统一的上层接口操作实验平台、进行大规模的网络实验。一方面考虑到图形化界面操作直观、实时性高、可交互性强，另一方面考虑到编程接口较高的灵活性和较高的可定制能力，所以 Klonet 的应用接口由图形接口和编

程接口共同组成, 不仅满足应用接口的易用性和功能性, 也能够应对大规模创新环境带来的挑战。

4.2 模拟网络构建

Klonet 构建的模拟网络, 是实际部署网络创新技术、开展创新实验的网络环境。为了提供真实、丰富的网络实验场景, 支持多种网络技术的行为验证和性能评估, 模拟网络应该具备任意拓扑的模拟能力, 多种网元功能、网络层次的模拟能力, 以及网络控制面和数据面的模拟能力。

第一, 为了实现任意拓扑的模拟能力, Klonet 首先构建一个具有任意拓扑的“二层网络”。从 overlay 网络的视角来看, 虚拟网元要按照任意的连接关系, 通过一条虚拟的二层链路互连起来, 虚拟网元的网络命名空间独立, 等同于物理的网络设备, 而虚拟链路要等同于由两个物理网络端口和一条物理网线形成的点到点连接, 从而在 overlay 网络的层面达成真正意义上的任意拓扑。第二, 实现多种网元功能、网络层次, 以及网络控制面的模拟能力, 实际是要将各层、各类的网络设备软件以及终端的网络应用安装、运行到“二层网络”的虚拟网元中, 而网络控制面则会由这些设备软件间的网络行为形成。第三, 为了实现网络数据面的模拟能力, Klonet 在“二层网络”中的端口设备上设置特定的端口规则, 模拟带宽、时延等数据面属性, 最终使该“二层网络”逐步构建为具有多种网络模拟能力的模拟网络, 如图 2 所示。

对于虚拟网元的实现。考虑到虚拟网元要具有独立的网络命名空间、要能够安装和运行软件应用, 并且要具备一定的性能隔离能力, 所以 Klonet 采用虚拟机或容器作为解决方案。虚拟机具有完整的操作系统内核, 性能隔离性好, 但较为笨重; 容器资源开销小、轻量且灵活, 但缺点在于容器共享宿主机内核, 性能干扰较大。因此要根据不同的性能需求来选择不同的虚拟化技术。

对于虚拟链路的实现。考虑到要构建等同于真实设备间的点到点连接, 网元中要有真实的网卡设备来配置队列规则, 所以对于同宿主机上的网元, Klonet 采用 vethpair^[10]以及 tap 与虚拟网桥的组合建立容器以及虚拟机之间的虚拟链路; 对于不同宿主机上的网元, Klonet 以前者虚拟链路为基础, 通过添加 VxLAN 隧道或配置 VLAN 实现虚拟链路。最后使用 Linuxtc 设置链路的数据面属性。

4.3 模拟网络管理

Klonet 支持在实验运行过程中对其模拟网络进行多种管理操作行为, 如网元的创建和删除、链路的增加和移除、创新实例的启动和停止等。这些操作行为不仅可以帮助用户将动态事件引入模拟网络, 使其具有真实场景下复杂多样的变化, 还可以帮助用户灵活地管理实例的运行状态、修改网络拓扑、改善实验方案。另外 Klonet 还提供

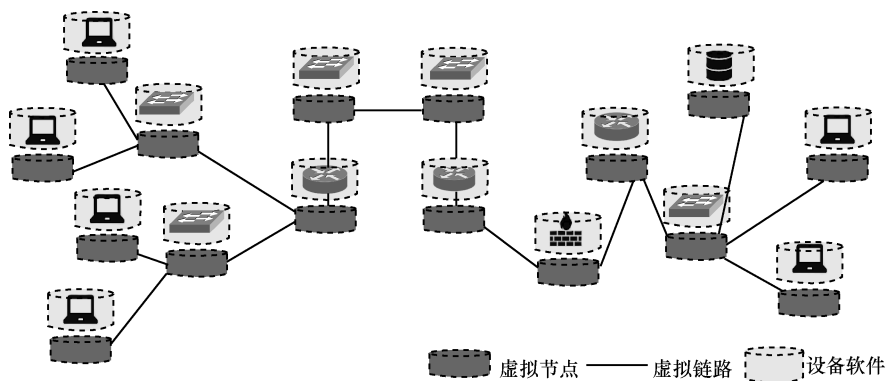


图 2 模拟网络构建方法



实时的网络拓扑展示功能，从而提高模拟网络的可视化水平和可操作性。

实现网络管理操作以及网络拓扑展示的关键在于模拟网络的描述信息，如果有数据模型能够维护模拟网络的状态信息，平台就可以根据这些数据进行相应的网元和链路操作，或根据数据模型来绘制图形化的拓扑界面。

对于模拟网络的展示，需要一种静态的模拟网络模型向图形接口描述当前时刻的网络状态以绘制对应的界面。当然，这类模型同样可以描述用户设计的网络拓扑，服务于网络拓扑定制功能。为此 Klonet 设计了 JSON 数据格式的拓扑描述文件，作为平台描述模拟网络的规范模型。但是由于其静态性，该模型很难跟踪模拟网络的动态改变、描述模拟网络的实时状态，而实验过程中对已部署网络的动态操作和修改又势必会改变其状态以及相应的模型表示。

所以 Klonet 还基于 Redis 设计了模拟网络的动态模型。动态模型的特点是会随着模拟网络状态的变化而变化。当短时间内有大量动态事件发生时，动态模型会随之产生大量更新，所以适合使用非关系的内存型数据库维护和存储。Redis 执行速度快、适用于频繁读写的场景，并且其键值（key-value）存储系统的性质，非常利于其动态表项与 JSON 格式的静态文件进行相互转换。综上，Klonet 采用 Redis 作为运维数据库，并以 Redis 的表项对模拟网络中的各类实体进行了重新刻画，最终形成网络的动态模型。

模拟网络部署时，静态的拓扑描述被实例化为虚拟网元和虚拟链路，同时也被实例化为运维数据库中的动态模型。当用户管理和操作模拟网络时，控制框架会以动态模型作为数据支持来执行相应的动作，并按照动作的执行结果进行相应的模型更新。对于拓扑展示，图形接口会定期向控制框架请求当前时刻的模拟网络描述，控制框架收到请求后将读取运维数据库中此刻的网络数

据，并把这些数据转换为网络的静态模型，然后返回到图形接口用于界面绘制。

4.4 实验支撑功能

考虑到拓扑结构、网络流量和实验结果等网络实验的基本要素^[9]，Klonet 对实验的支撑能力主要包括网络拓扑定制、多种流量发生和实验结果分析的功能，给予网络实验必要的功能助力，节省开发、测试、分析网络创新技术所需的时间，简化实验过程，降低技术创新门槛。在实验中，网络拓扑定制实现任意网络拓扑的灵活设计和部署，多种流量发生向模拟网络注入不同分布、不同模式的流量，实验结果分析提供对网络指标和协议行为的监控、统计、分析和呈现，如图 3 所示。

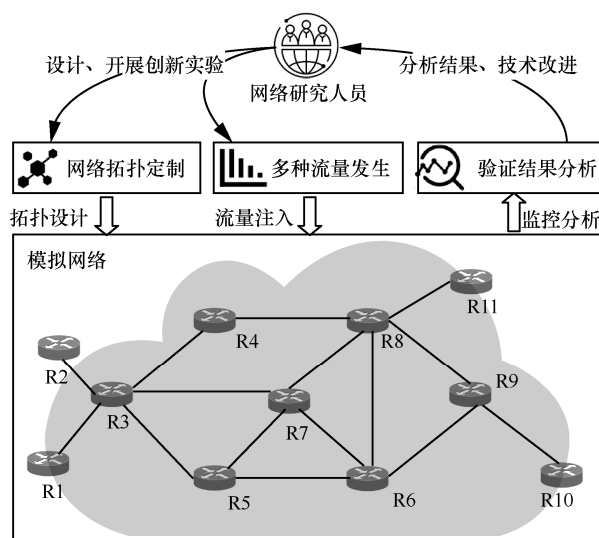


图 3 Klonet 实验支撑功能

(1) 网络拓扑定制

Klonet 提供两类拓扑定制方法，图形化接口定制和编写拓扑描述文件定制，分别对应使用图形接口和编程接口两种部署方式。图形化方法在采用典型拓扑模板（如 Tree、FatTree、BCube 等）的基础上手动拖拽网元、链路的形式定制拓扑；或者可以完全由手动拖拽网元、链路设计。典型模板用来快速生成大规模的拓扑骨干，手动拖拽则在此基础上进行微调。拓扑设计完成之后，图

形接口会生成相应的规范化拓扑描述文件。该方法可有效减轻拓扑结构设计负担,适应特异网络拓扑结构的交互式设计要求,使拓扑定制更加直观、便捷,交互性更强。编写拓扑描述文件的方法要求实验人员向平台提交可读的拓扑描述文件,描绘定制后的拓扑结构。所以相较于交互式的手动拖曳,通过程序脚本可以自动化生成规范的拓扑描述,能够高效定制大规模、有规律的拓扑,并且可以解析已有的拓扑数据集和其他平台的拓扑文件,便于现有网络描述的移植,丰富定制拓扑结构的方法。

(2) 多种流量发生

定制好的网络拓扑被部署到硬件集群后,若没有业务流量的注入,则只存在路由交换协议、DHCP 等控制平面协议产生的控制面流量,无法满足对数据面操作行为的观察和性能指标的评估。为了解决这一问题,Klonet 提供包含数据流生成和数据包生成的多种流量发生功能,向网络数据面注入可灵活定制的高速数据流量。数据流生成基于系统的传输层接口,Klonet 允许用户定制流请求大小的累积分布函数,请求总时间和期望吞吐等参数。数据包生成则凭借构包、发包的方式产生流量,Klonet 允许用户定制报文的字段格式、取值和长度,允许定制包长的分布,允许按照符合指定分布的时间间隔发送数据包。另外,Klonet 还将 Iperf 以程序化接口的形式提供给用户,用户可以利用 Iperf 产生背景流,也可以在 Iperf 的基础上通过编程的方式灵活定义流量发生模式。

(3) 实验结果分析

实验结果分析是开发者在平台上开发、验证网络创新技术的最后一步,为此 Klonet 的实验结果分析功能从控制面的协议行为分析和数据面的性能结果获取两个方向出发,对模拟网络进行监测。网络监测主要基于网络端口上的报文抓取,为了提高抓包效率、减小抓包开销,平台引入可

编程的 eBPF 技术在内核中设置钩子捕获事件信息,使报文信息获取更加轻量且赋予其更加丰富的定制化能力。对于可定制的协议分析部分,平台提供对常用协议的报文过滤能力以及字段提取、分析能力,可按照协议类型或字段类型对报文进行抓取、聚合等操作。对于性能结果的获取,平台提供部分常见性能指标(如吞吐率)的计算获取。与网元资源用量有关的分析和验证,平台主要通过集成 Prometheus 监控系统实现,支持对计算、存储等资源使用情况的收集、存储和处理。实验结果分析功能允许通过可视化接口或描述文件对具体的监控事件进行定制,支持监测数据的可视化分析,并且允许以文件形式导出原始数据和结果图像。

5 平台用例

通过对两类面向不同网络研究问题的实验场景分析,具体说明平台对网络技术创新的支撑能力,以及利用平台开展实验的大致方法。

5.1 端到端拥塞控制

TCP 作为互联网的主流传输协议,其拥塞控制方法对端到端数据传输的性能表现和带宽资源的使用情况具有重大影响。随着近年来网络环境愈发复杂、网络应用场景愈发多样的发展趋势,网络拥塞控制方法仍在面临着源源不断的需求和挑战,该研究也一直都是网络创新中的热点。

为了在现实中评估拥塞控制算法,基于 UDP 套接字的用户态实现方式通常会被采用^[11],该方法不用改动内核协议栈中的具体实现,可以作为用户态应用程序执行。另外,为了理解算法的真实表现,通常要在广域网链路、卫星链路等不同的网络条件下对其进行测试。

在用户态进行拥塞控制研究以及模拟复杂的网络链路场景 Klonet 都能够支持。首先,Klonet 构建的模拟网络,其虚拟网元采用虚拟机或容器实现,具有完整的内核协议栈,易于封装和执行用户态的算法实现。其次,Klonet 能够模拟网络



数据面的多种属性并支持属性的定制化配置，具体见表 1。通过对带宽、往返时延、丢包率以及端口缓存大小等属性的有机组合，可以模拟出多种不同的网络条件用于性能表现的评估（如吞吐率、缓存占用率、时延敏感性、公平性等）。最后，Klonet 的 Web 接口可以帮助用户进行可视化的拓扑定制和链路属性配置，在实验过程中便捷引入新的链路竞争以及改变网络属性。Klonet 使得拥塞控制算法的开发和实验可以更灵活、高效地开展。

表 1 网络属性定制能力

网络属性	定制能力
带宽 (bandwidth)	固定值或在区间内随
往返时延 (round-trip time, RTT)	机变化
丢包率 (loss rate)	
端口缓存 (buffer)	

5.2 IPv6 段路由协议

随着 5G 时代的到来和网络应用场景的不断扩展，基于 IPv6 的网络应用创新和网络协议创新成为下一代互联网的重要发展动力。其中，IPv6 段路由 SRv6 作为最具代表性的创新协议之一，凭借其易扩展、易编程、高可靠等特性吸引了广大的研究人员，也有不少方案已经落地部署。

利用 Klonet 可以设计、搭建完整的网络测试环境，用于 SRv6 创新实例的功能性验证，参考网络场景如图 4 所示。该网络由两部分组成，一部分是控制和管理平面，另一部分是数据平面。对于数据平面的模拟，Klonet 能够在“二层网络”的虚拟节点中安装、运行 FRR (FRRouting) 等软件路由协议族使其成为基于内核的 SRv6 路由器网元。通过 eBPF 和 XDP 等可编程技术，可以在内核层面对路由器网元进行功能扩展和数据路加速，实现 SRv6 网络的数据平面创新。对于控制平面的模拟，Klonet 单独为网元和控制器创建一个上层的控制网络，用于 SRv6 南向接口的创新实现，并且当路由网络的测试场景规模较大时，可以利用 Klonet 编

程接口在可定制程度和大批量操作等方面的优势，采用多样的方式编写拓扑描述文件、定制网络拓扑，成批地引入各类网络事件、执行各类操作。

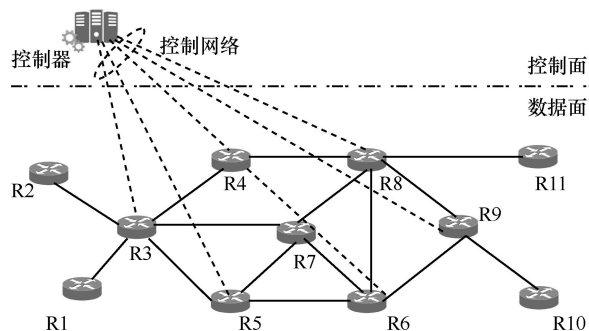


图 4 SRv6 网络场景参考

6 结束语

本文面向网络技术创新在开发和实验方面的需求，提出了网络模拟实验平台 Klonet 的设计目标和具体实现方法。Klonet 首先能够满足大规模的实验场景，并且其搭建的模拟网络具有任意拓扑、在功能和属性上十分贴近真实的网络情况，使多类别的网络技术创新能够部署、运行到多样化的网络场景中，以进行相应的正确性验证和性能表现评估，最后凭借实验过程中的拓扑展示能力、网络操作能力，以及对实验从拓扑到流量再到结果的全方位功能支持，带来更多方案上的可行性和操作上的便利性。Klonet 能够应对由网络规模不断扩大和技术创新形式不断丰富所带来的挑战，有望助力网络技术的创新和发展。

现有 Klonet 是基于容器的虚拟化技术实现，更适合进行功能方面的创新验证，后续将考虑采用高速互联的多卡高性能服务器集群来支撑性能方面的创新试验。同时，现有 Klonet 对模拟网络的优化部署研究不够深入，导致平台资源利用率不高。未来 Klonet 将会关注平台各类资源的量化描述，通过大量测试建立合理的资源需求模型，并以此为基础设计模拟网络切分和放置算法，从而实现资源的高效利用。

参考文献:

- [1] BANKS J. Discrete event system simulation[EB]. 2005.
- [2] RILEY G F, HENDERSON T R. The ns-3 network simulator[M]//Modeling and Tools for Network Simulation. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010: 15-34.
- [3] CHANG X J. Network simulations with OPNET[C]//Proceedings of WSC'99.1999 Winter Simulation Conference Proceedings. 'Simulation - A Bridge to the Future' (Cat. No.99CH37038). Piscataway: IEEE Press, 1999: 307-314.
- [4] CHUN B, CULLER D, ROSCOE T, et al. PlanetLab: an overlay testbed for broad-coverage services [J]. ACM SIGCOMM Computer Communication Review, 2003, 33(3): 3-12.
- [5] Stoller M H R R L, Duerig J, Guruprasad S, et al. Large-scale virtualization in the emulab network testbed[C]//USENIX 2008 Annual Technical Conference. Berkeley: USENIX Association, 2008: 255-270.
- [6] LANTZ B, HELLER B, MCKEOWN N. A network in a laptop: rapid prototyping for software-defined networks[C]// Proceedings of the Ninth ACM SIGCOMM Workshop on Hot Topics in Networks - Hotnets '10. New York: ACM Press, 2010: 1-6.
- [7] 蒋林涛. 数据网的现状及发展方向[J]. 电信科学, 2019, 35(8): 2-15.
JIANG L T. Current situation and development trend of data network[J]. Telecommunications Science, 2019, 35(8): 2-15.
- [8] LIU H H, ZHU Y B, PADHYE J, et al. CrystalNet: faithfully emulating large production networks[C]//Proceedings of the 26th Symposium on Operating Systems Principles. Shanghai China. New York: ACM Press, 2017: 599-613.
- [9] SAINO L, COCORO C, PAVLOU G. A toolchain for simplifying network simulation setup[C]//Proceedings of the Sixth International Conference on Simulation Tools and Techniques. New York: ACM Press, 2013: 82-91.
- [10] HANDIGOL N, HELLER B, JEYAKUMAR V, et al. Reproducible network experiments using container-based emulation[C]//Proceedings of the 8th international conference on Emerging networking experiments and technologies - CoNEXT '12. New York: ACM Press, 2012: 253-264.
- [11] DONG M, MENG T, ZARCHY D, et al. PCC vivace: on-line-learning congestion control[C]//15th USENIX Symposium on Networked Systems Design and Implementation (NSDI 18). Berkeley: USENIX Association, 2018: 343-356.

[作者简介]



谢景昭(1996-), 男, 电子科技大学硕士生, 主要研究方向为网络功能虚拟化、云计算、软件定义网络等。



单炜(1998-), 男, 电子科技大学硕士生, 主要研究方向为网络功能虚拟化、云计算、软件定义网络等。



肖畅(1997-), 男, 电子科技大学硕士生, 主要研究方向为网络功能虚拟化、云计算、软件定义网络等。



马铁(1998-), 男, 电子科技大学硕士生, 主要研究方向为网络功能虚拟化、云计算、软件定义网络等。

陈力(1989-), 男, 博士, 华为技术有限公司系统理论研究员, 主要研究方向为人工智能和网络交叉的方向。

虞红芳(1975-), 女, 电子科技大学教授, 主要研究方向为下一代网络、分布式系统、网络安全、人工智能等。

孙罡(1979-), 男, 电子科技大学教授, 主要研究方向为网络虚拟化、分布式系统、网络安全等。